REMARKS

This communication is in response to the Final Office Action mailed on July 13, 2007. In the Final Office Action, claims 17-21, 23-26, and 27-30 were pending. Claims 17-21 and 23-26 have been cancelled and new claims 31-33 have been added. Thus, claims 27-33 remain pending.

CLAIM REJECTIONS - 35 U.S.C. § 112

Claims 17-21, and 23-26 were rejected under 35 U.S.C. § 112 as failing to comply with the written description requirement. These claims have been cancelled and thus withdrawal of this rejection is requested.

CLAIM REJECTIONS - 35 U.S.C. § 102

Claims 27-30 were rejected under 35 U.S.C. § 102(e) as being anticipated by Wu. Claim 27 and new claim 31 are independent. These claims are directed to providing evaluation of models for particular word types. For each individual type of word, the effectiveness can be measured, which provides a useful tool for evaluating word segmentation models and gives indications for ways of improvement. Wu describes customizable segmentation of morphologically derived words in Chinese. In section 2, Wu identifies guidelines for segmenting Chinese text, given that there are several different criteria to be used in Chinese word segmentation. Several different situations from Chinese text can be recognized for use in segmentation, such as reduplication, affixation, compounding, merging, splitting, named entities, and factoids. Using the guidelines, a language model can be developed for annotating a corpus. Section 3.3 discusses evaluation of the annotations for word segmentation models built using the guidelines and several adjustable parameters. The evaluation is performed for each model as a whole and simply does not discuss evaluation based on a particular word type or providing effectiveness indications for particular word types.

By focusing on individual word types, an improved evaluation approach is realized that can be utilized for determining effectiveness of word segmentation models. The effectiveness can be evaluated to give a better indication of how well different models are able to segment text and/or

how well types of words can be recognized.

Independent claim 27 has been amended to recite a method for evaluating word segmentation language models. The method includes building the word segmentation model based on an annotated corpus and applying the language model to a test corpus of unsegmented text different from the annotated corpus to provide an output indicative of words in the test corpus and a word type indication for each word. The word type indication is one of a plurality of word type indications. The method also includes comparing the word type indication for each word in the output of the language model with the predefined word type indications of words of the test corpus. The language model is evaluated based on the comparison of the word type indication for each word in the output and the predefined word type indications to provide an indication of effectiveness of the language model as a function of the word type indications identified by the language model.

New independent claim 31 recites a method of evaluating word segmentation models. The method includes using a first word segmentation model to segment a corpus of text into words and apply tags to the words indicative of one of the plurality of word types. The words and tags form a first output. A second word segmentation model is used to segment the corpus of text into words and apply tags to the words indicative of one of the plurality of word types. The words and tags form a second output. The first output is compared to a predefined indication of words and tags of the words indicative of one of the plurality of word types from a corpus of text to provide a first set of values for each of the plurality of word types. The first set of values is indicative of how the first word segmentation model recognizes each of the plurality of word types. The second output is compared to the predefined indication of words and tags of the words indicative of one of the plurality of word types from a corpus of text to provide a second set of values for each of the plurality of word types. The second set of values is indicative of how the second word segmentation model recognizes each of the plurality of word types. The first set of values and the second set of values is compared to determine effectiveness of the first word segmentation model and the second word segmentation model with respect to each of the plurality of word types.

When compared to the evaluation described in section 3.3 of Wu, both independent claims

27 and 31 provide evaluation methods for word segmentation models that are simply not taught or suggested by that of Wu. In contrast to Wu, the evaluation method recited in claims 27 and 31 can be utilized to provide a more precise evaluation in the development of word segmentation models. As a result, word segmentation models can be evaluated in a different manner to achieve effective evaluation.

In view of the foregoing, Applicants submit that the present application is in condition for allowance. Favorable action on all claims is respectfully requested.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____

Todd R. Fronek, Reg. No. 48,516
900 Second Avenue South, Suite 1400
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

TRF:lah